

新興国のテクニカルターム検証方法

タイ語・ベトナム語・インドネシア語の
テクニカルターム収集

佐野浩太郎(株式会社IHI)
○武藤亜弓(株式会社ネットス)

発表の流れ

1.はじめに

1-1.ASEAN諸国と日本

1-2.ASEAN諸国の特許調査が困難な理由

1-3.本研究の目的

2.3つのアプローチ方法

2-1.3つのアプローチ方法について

2-2-1.言語からのアプローチについて

2-2-2.言語からのアプローチによる収集結果

2-3-1.分類からのアプローチについて

2-3-2.分類からのアプローチによる収集結果

2-4-1.ファミリーからのアプローチについて

2-4-2.ファミリーからのアプローチによる収集結果

2-5.3つのアプローチ方法による収集結果

3.まとめ

1.はじめに

1-1.ASEAN諸国と日本

1-2.ASEAN諸国の特許調査が困難な理由

1-3.本研究の目的

2.3つのアプローチ方法

2-1.3つのアプローチ方法について

2-2-1.言語からのアプローチについて

2-2-2.言語からのアプローチによる収集結果

2-3-1.分類からのアプローチについて

2-3-2.分類からのアプローチによる収集結果

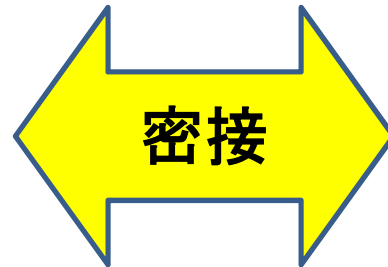
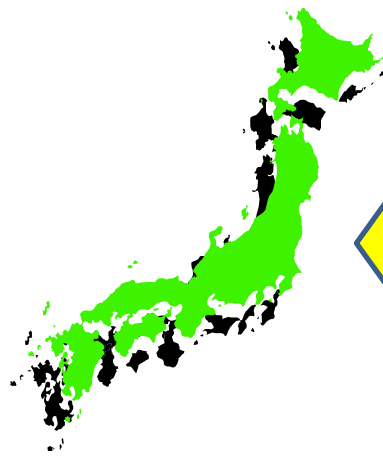
2-4-1.ファミリーからのアプローチについて

2-4-2.ファミリーからのアプローチによる収集結果

2-5.3つのアプローチ方法による収集結果

3.まとめ

1-1.ASEAN諸国と日本



国境を越えた
生産ネット
ワーク

日中関係の
影響



2013年
日本企業投資額: 過去最高
日本企業からの注目度【高】

BUT !!

**新興国における特許
調査は困難！！**

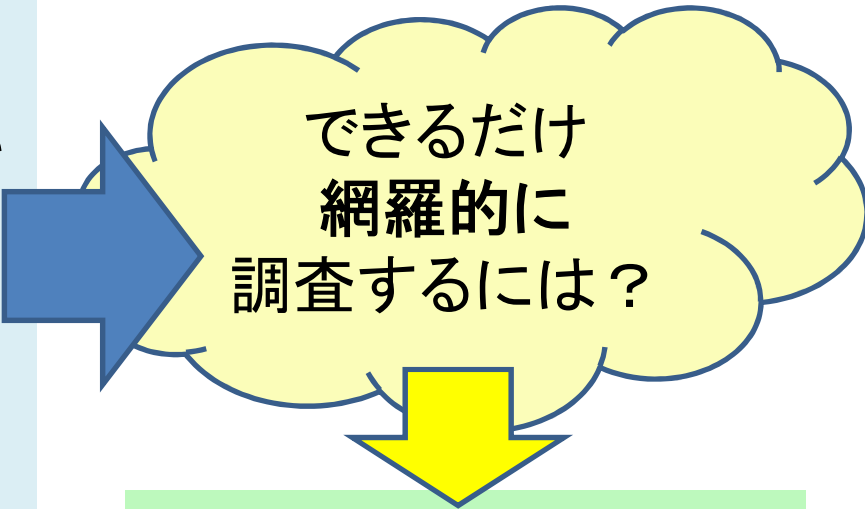
1-2.ASEAN諸国の特許調査が困難な理由

①特許情報の未整備

- ・法制度が整っていない
- ・データベースが整っていない
- ・特許分類の付与が不十分

②言語の障壁

- ・特許公報の言語が、馴染みのある言語(英語・中国語等)ではない場合が多い



できるだけ
網羅的に
調査するには？

キーワード検索が
欠かせない！！

1-3.本研究の目的

- IPC付与が不十分であり、英語収録がされていない新興国で、テクニカルタームがどのように使われているのか検証する。
- 新興国においてテクニカルタームを網羅的且つ効率的に収集できる方法を体系化し、提案する。

検証方法:

「半導体」という言葉について、タイ語・ベトナム語・インドネシア語におけるテクニカルタームを調べる。

1.はじめに

1-1.ASEAN諸国と日本

1-2.ASEAN諸国の特許調査が困難な理由

1-3.本研究の目的

2.3つのアプローチ方法

2-1.3つのアプローチ方法について

2-2-1.言語からのアプローチについて

2-2-2.言語からのアプローチによる収集結果

2-3-1.分類からのアプローチについて

2-3-2.分類からのアプローチによる収集結果

2-4-1.ファミリーからのアプローチについて

2-4-2.ファミリーからのアプローチによる収集結果

2-5.3つのアプローチ方法による収集結果

3.まとめ

2-1.3つのアプローチ方法について

3つのアプローチ

- (1) 言語からのアプローチ
- (2) 分類からのアプローチ
- (3) ファミリーからのアプローチ



特許分類付与が不十分
⇒(1)～(3)の方法を組み合わせる必要性あり

2-2-1. 言語からのアプローチについて

- ①-1 日本語→機械翻訳で対象言語に変換
- ①-2 英語→機械翻訳で対象言語に変換
- ② 各国特許庁DBまたは原語収録している商用DB
でキーワード検索(タイトル+要約)
- ③ ヒット公報の内容を読む(機械翻訳で変換)

- (-) 機械翻訳で変換される言い方以外の言葉に気付きにくい
- (-) 機械翻訳の正確性に左右される
- (-) データベースの仕様によって結果が変わる
- (+) 該当分類(IPC)調査の参考になる

2-2-2. 言語からのアプローチによる収集結果

日本語・英語	タイ語	公報ヒット 件数
半導体	สารกึ่งตัวนำ	776
semiconductor	สารกึ่งตัวนำ	776
	ครึ่งสื่อไฟฟ้า	0
	เซมิคอนดัก	49
	เซมิคอนดักเตอร์	14
	เซมิคอนดักเตอร์	46

日本語・英語	ベトナム語	公報ヒット 件数
半導体	bán dẫn	62
semiconductor		

日本語・英語	インドネシア語	公報ヒット 件数
半導体	semikonduktor	126
semiconductor		

- ・タイ語「ครึ่งสื่อไฟฟ้า」は公報ヒット件数0件
- ・ベトナム語、インドネシア語は共に1通り

2-3-1.分類からのアプローチについて

- ① 対象国特許庁DB、または対象国対応の商用DBで分類検索
- ② ヒット公報の内容を読む(機械翻訳で変換)

- (-)分類を正確に付与されている公報が少ない
(付与自体されていない公報もあり)
- (-)上位の分類で検索すると、ノイズが増えてしまい検証しにくい
- (+)ノイズがあるとはいえ、比較的近い技術分野のものが拾える
- (+)「言語からのアプローチ」では見つけられなかった言葉が見つかる可能性がある

2-3-2.分類からのアプローチによる収集結果

日本語	タイ語	件数 ※
半導体	กึ่งตัวนำ	25
	สารกึ่ง	4
	เซมิ	1
	คอนดัคเตอร์	1
半導体デバイス (記憶装置+半導体)	อุปกรณ์สารกึ่งตัวนำ	
半導体デバイス	อุปกรณ์กึ่งตัวนำ	
半導体ウエハ	เวเฟอร์กึ่งตัวนำ	
	เวเฟอร์สารกึ่ง	
半導体チップ	ชิพสารกึ่งตัวนำ	
発光ダイオード	ไดโอดเปล่งแสง	
半導体基板	ฉับสเตรทกึ่งตัวนำ	

※件数とは、「H01L21」172件を閲覧した中で、使用されていた公報の件数

日本語	ベトナム語	件数 ※
半導体	bán dẫn (BÁN DẪN)	7

※件数とは、「H01L21」35件を閲覧した中で、使用されていた公報の件数

日本語	インドネシア語	件数 ※
半導体	semikonduktor	7
半導体基板	substrat semikonduktor	
半導体膜	selaput semikonduktor	
半導体フィルム	film semikonduktor	

※件数とは、「H01L21」105件を閲覧した中で、使用されていた公報の件数

- ・タイ語はH01L21の172件中31件
- ・ベトナム語はH01L21の35件中7件
- ・インドネシア語はH01L21の105件中7件

2-4-1.ファミリーからのアプローチについて

- ①対象国を収録し、且つファミリーを検索できる特許データベースを使用し、国コードを対象国に指定、該当IPCを入力する
- ②ヒット公報のファミリーのうち対象国の公報番号を抽出
- ③各国特許庁DBまたは原語収録している商用DBで、番号を検索、ヒット公報の内容を読み、該当する言葉を抽出
(機械翻訳で変換)

- (-) 公報番号の表記方法がデータベース毎に異なることがある
- (-) 特許庁DBで1件ずつ公報を検索するため時間がかかる
- (-) 国や商用データベースによって、ファミリーの収録率が異なる
- (+) 求めている技術分野がヒットしやすく、ノイズが少ない

2-3-2.分類からのアプローチによる収集結果

日本語	タイ語	公報ヒット件数
半導体	การเซมิคอน	0
	อุปกรณ์ตัวนำ	0
半導体デバイス	อุปกรณ์กึ่งตัว	15
	อุปกรณ์สารกึ่งตัวนำ	228
ウエハー	เวเฟอร์	
	กายเวเฟอร์	
	ซองเวเฟอร์	
	กับแผ่นเวเฟอร์	
	แผ่นเวเฟอร์	
ウェーハ	เวเฟอร์แค	

日本語	ベトナム語	公報ヒット件数
半導体	bán dẫn	62

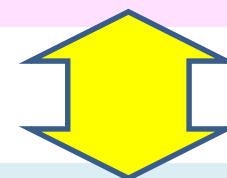
日本語	インドネシア語	公報ヒット件数
半導体	semikonduktor	126
	semi konduktor	18
	semi-konduktor	10

- ・収集した言葉を再度各国特許庁データベースで検索してもヒットしない場合あり(特許庁DBのキーワード検索機能に問題がある可能性)
- ・(1)や(2)のアプローチ方法では見つからない言葉を収集できる

2-5.3つのアプローチ方法による収集結果

日本語	タイ語・ベトナム語・インドネシア語	公報ヒット件数	抽出アプローチ方法			
			①	②	③	
半導体	タイ語	สารกึ่งตัวนำ	776	○	○	○
		ครึ่งลือไฟฟ้า	0	○		
		เซมิคอนดัก	49	○		○
		เซมิคอนดักเตอร์ที่	14	○		○
		เซมิคอนดักเตอร์	46	○		○
		กึ่งตัวนำ	1077		○	○
		สารกึ่ง	866		○	○
		เซมิ	275		○	○
		คอนดักเตอร์	55		○	○
		การเซมิคอน	0			○
半導体デバイス	タイ語	อุปกรณ์สารกึ่งตัวนำ	228		○	○
		อุปกรณ์กึ่งตัวนำ	105		○	
		อุปกรณ์กึ่งตัว	15			○
		อุปกรณ์งตัวนำ	0			○
半導体	ベトナム語	bán dẫn	62	○	○	○
半導体	インドネシア語	semikonduktor	126	○	○	○
		semi konduktor	18			○
		semi-konduktor	10			○

3つのアプローチ方法すべてで抽出できるものは、使用頻度が高い（一般的な言葉と推察）



ヒット件数の多い言葉が、どのアプローチ方法でも抽出できるとは限らない

2-5.3つのアプローチ方法による収集結果

日本語	タイ語・ベトナム語・インドネシア語	公報ヒット件数	抽出アプローチ方法			
			①	②	③	
半導体	タイ語	สารกึ่งตัวนำ	776	○	○	○
		เครื่องสี่ไฟฟ้า	0	○		
		เซมิคอนดัก	49	○		○
		เซมิคอนดักเตอร์ที่	14	○		○
		เซมิคอนดักเตอร์	46	○		○
		กึ่งตัวนำ	1077		○	○
		สารกึ่ง	866		○	○
		เซมิ	275		○	○
		คอนดักเตอร์	55		○	○
		การเซมิคอน	0			○
半導体デバイス	タイ語	อุปกรณ์สารกึ่งตัวนำ	228		○	○
		อุปกรณ์กึ่งตัวนำ	105		○	
		อุปกรณ์กึ่งตัว	15			○
		อุปกรณ์กึ่งตัวนำ	0			○
半導体	ベトナム語	bán dẫn	62	○	○	○
半導体	インドネシア語	semikonduktor	126	○	○	○
		semi konduktor	18			○
		semi-konduktor	10			○

表記パターンが多いが、
表記揺れ等の問題から、
更に増える可能性あり

公報収録の問題(文字
化け等)から、この表記
以外にも存在する可能
性あり

複合語は、「-」や空白が
ある等、複数の表記パ
ターンあり

1.はじめに

1-1.ASEAN諸国と日本

1-2.ASEAN諸国の特許調査が困難な理由

1-3.本研究の目的

2.3つのアプローチ方法

2-1.3つのアプローチ方法について

2-2-1.言語からのアプローチについて

2-2-2.言語からのアプローチによる収集結果

2-3-1.分類からのアプローチについて

2-3-2.分類からのアプローチによる収集結果

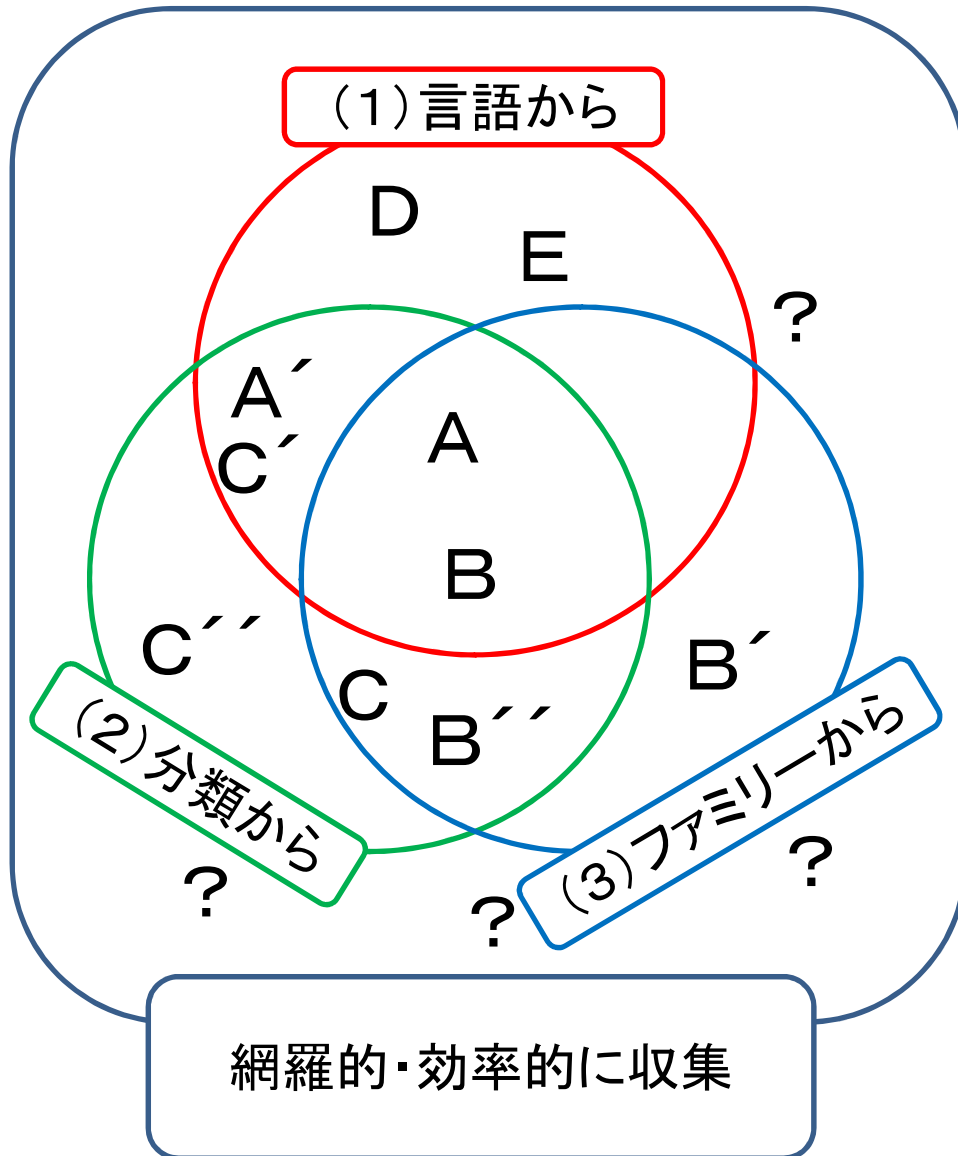
2-4-1.ファミリーからのアプローチについて

2-4-2.ファミリーからのアプローチによる収集結果

2-5.3つのアプローチ方法による収集結果

3.まとめ

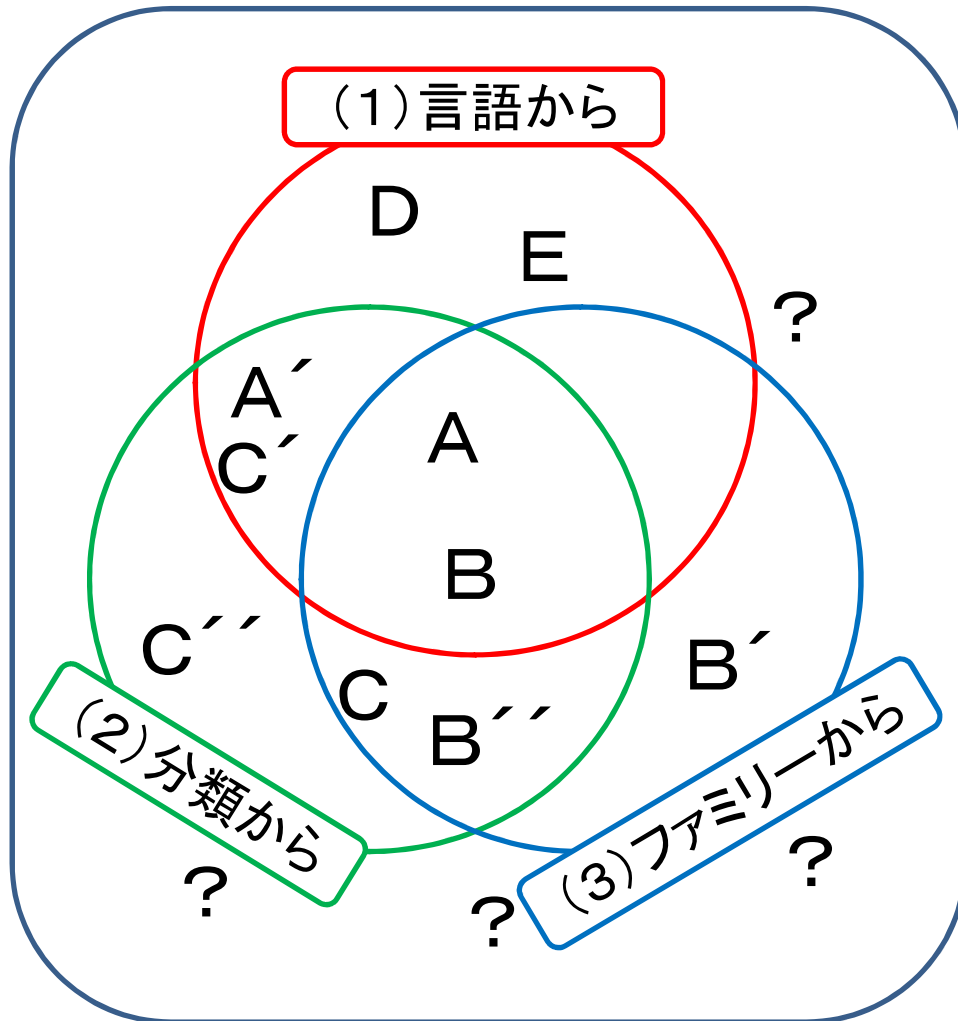
3.まとめ



網羅的: 3つのアプローチを組み合わせることで、なるべく漏れを少なくする

効率的: 3つのアプローチを順番に行っていく。やみくもに収集しない。

3.まとめ



(1)言語から

どの機械翻訳・辞書ツール
を使用するか

(2)分類から

どのIPCを使用するか

(3)ファミリーから

どの商用データベース(ファ
ミリー)を使用するか

(留意点)3つの選択により、その網羅性や作業効率には変化が生じる

謝辞

本報告は2014年度の「アジア特許情報研究会」のワーキングの一環として報告するものであり、研究会の皆様、特に、故太細博利氏に様々な協力をしていただきましたこと、改めて感謝申し上げます。

ご清聴ありがとうございました